

  Nucleotide

PubMed Nucleotide Protein Genome Structure PMC Taxonomy OMIM Boo

Search for

Limits Preview/Index History Clipboard Details

Show:

☐ 1: NM_173674. Homo sapiens disc...[gi:27735142]

Links

LOCUS NM_173674 2010 bp mRNA linear PRI 05-OCT-2003
DEFINITION Homo sapiens discoidin, CUB and LCCL domain containing 1 (DCBLD1), mRNA.
ACCESSION NM_173674
VERSION NM_173674.1 GI:27735142
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2010)
AUTHORS Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D., Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S., Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Helton,E., Kettelman,M., Madan,A., Rodrigues,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E., Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
TITLE Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
MEDLINE 22388257
PUBMED 12477932
COMMENT PROVISIONAL REFSEQ: This record has not yet been subject to final NCBI review. The reference sequence was derived from BC035671.1.
FEATURES
Location/Qualifiers
source 1..2010
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="6"
/map="6q22.31"
gene 1..2010
/gene="DCBLD1"
/note="synonyms: MGC46341, dJ94G16.1"
/db_xref="GeneID:285761"

CDS /db_xref="LocusID:285761"
126..1745
/gene="DCBLD1"
/note="go_process: cell adhesion [goid 0007155] [evidence IEA]"
/codon_start=1
/product="hypothetical protein MGC46341"
/protein_id="NP_775945.1"
/db_xref="GI:27735143"
/db_xref="GeneID:285761"
/db_xref="LocusID:285761"
/translation="MVPGARGGGALARAAGRLLALLLAVSAPLRLQAEELGDGCGHL
VTYQDSGMTSKNYPGTYPNHTVCEKTIITVPKGKRLILRLGDLIDIESQTCASDYLLFT
SSSDQYGPYCGSMTVPKELLNTSEVTVRFESGSHISGRGFLITYASSDHPDLITCLE
RASHYLKTEYSKFCPAGCRDVAGDISGNMVDGYRDTSLCKAAIHAGIIADELGGQIS
VLQRKGISRYEGILANGVLSRDGSLSDKRFLFTSNGCSRSLSFEPDQGIRASSSWQSV
NESGDQVHWSPGQARLQDQGPSWASGDSSNNHKPREWLEIDLGEKKKITGIRTTGSTQ
SNFNFYVKSFVMNFKNNNSKWKTYKGIVNNEEKVFQGNSNFRDPVQNNFIPPIVARYV
RVVPQTHQRIALKVELIGCQITQGNDLSLVWRKTSQSTSVSTKKEDETITRPIPEET
STGINITTVAIPLVLLVLFVAGMGIFAFAFRKKKKKGGSPYGSAAEQKTCWKQIKYPF
ARHQSAEFTISYDNEKEMTQKLDLITSDMAG"

misc_feature 246..572
/gene="DCBLD1"
/note="CUB; Region: Domain first found in C1r, C1s, uEGF,
and bone morphogenetic protein. This domain is found
mostly among developmentally-regulated proteins.
Spermadhesins contain only this domain"
/db_xref="CDD:smart00042"

misc_feature 591..869
/gene="DCBLD1"
/note="LCCL; Region: LCCL domain"
/db_xref="CDD:pfam03815"

misc_feature 912..1352
/gene="DCBLD1"
/note="F5_F8_type_C; Region: F5/8 type C domain. This
domain is also known as the discoidin (DS) domain family.
The bacterial examples are not yet included in the SEED
alignment and are only found with low scores"
/db_xref="CDD:pfam00754"

BASE COUNT 577 a 457 c 512 g 464 t
ORIGIN
1 ggcagcgact gcgccccgtc ccggcgccgc gctcgtcgcg agaggaggcg gcccgggccc
61 ggcagctgcg gctcgggatc cgtcagggg aggccgagct tgccaagctg gcgcccagcg
121 gggatcatggt gcccgggcgc gcgcgccgcg gcgcactggc gcgggctgcc gggcggggccc
181 tcctggcttt gctgctcgcg gtctccgccc cgctccggct gcaggcggag gagctgggtg
241 atggctgtgg acacctagt acttatcagg atagtggcac aatgacatct aagaattatc
301 ccgggacctc cccaatcac actgtttgcg aaaagacaat tacagtacca aaggggaaaa
361 gactgattct gaggttgga gatttgata tcgaatccca gacctgtgct tctgactatc
421 ttctcttcac cagctcttca gatcaatatg gtccatactg tggaagtatg actgttccca
481 aagaactctt gttgaacaca agtgaagtaa ccgtccgctt tgagagtgga tcccacattt
541 ctggccgggg ttttttgctg acctatgcga gcagcgacca tccagattta ataactgtt
601 tggaacgagc tagccattat ttgaagacag aatacagcaa attctgccc gctggttgta
661 gagacgtagc aggagacatt tctgggaata tggtagatgg atatagagat acctctttat
721 tgtgcaaagc tgccatccat gcaggaataa ttgctgatga actaggtggc cagatcagtg
781 tgcttcagcg caaagggatc agtcgatatg aagggtattt ggccaatggt gttctttcga
841 gggatgggtc cctgtcagac aagcgatttc tgtttacctc caatggttgc agcagatcct
901 tgagttttga acctgacggg caaatcagag cttcttcctc atggcagtcg gtcaatgaga
961 gtggagacca agttcactgg tctcctggcc aagcccgact tcaggaccaa ggcccatcat
1021 gggcttcggg cgacagtagc aacaaccaca aaccacgaga gtggctggag atcgatttgg
1081 gggagaaaaa gaaaataaca ggaattagga ccacaggatc tacacagtcg aacttcaact

1141 tttatgttaa gagttttgtg atgaacttca aaaacaataa ttctaagtgg aagacctata
1201 aaggaattgt gaataatgaa gaaaagggtg ttcagggtaa ctctaacttt cgggaccag
1261 tgcaaaacaa tttcatccct cccatcgtgg ccagatatgt gcgggttgtc cccagacat
1321 ggcaccagag gatagccttg aagggtggagc tcattggttg ccagattaca caaggtaatg
1381 attcattggt gtggcgcaag acaagtcaaa gcaccagtgt ttcaactaag aaagaagatg
1441 agacaatcac aaggcccatc ccctcggaag aaacatccac aggaataaac attacaacgg
1501 tggctattcc attggtgctc cttgttgctc tgggtgttgc tggaatgggg atctttgcag
1561 cctttagaaa gaagaagaag aaaggaagtc cgtatggatc agcagaggct cagaaaacag
1621 actgttgga gacagattaaa tatccctttg ccagacatca gtcagctgag tttaccatca
1681 gctatgataa tgagaaggag atgacacaaa agttagatct catcacaagt gatatggcag
1741 gttaactccg ttgactgcca aaatagcatc cccaacgtgc agccctccgc atctatcagc
1801 aggttgcccc ggatggatct cagagatgag gatcggaaca ccatgttctt tcccacccta
1861 acaacaacaa agggcagtaa attaaagtac tctttgtaag gtacagttac cgattaatct
1921 agagataaaa tattttctta aaaatatatt tcattaaaca cctatgctgt ctctataaaa
1981 aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa

//

[Disclaimer](#) | [Write to the Help Desk](#)
[NCBI](#) | [NLM](#) | [NIH](#)

Oct 2 2003 18:31:01

FASTA searches a protein or DNA sequence data bank
version 3.3t05 March 30, 2000

Please cite:

W.R. Pearson & D.J. Lipman PNAS (1988) 85:2444-2448

/tmp/fastaDAE.aqrX: 1464 nt
>LEX 71 Seq id no:1
vs /tmp/fastaEAAF.aqrX library
searching /tmp/fastaEAAF.aqrX library

2010 residues in 1 sequences

FASTA (3.34 January 2000) function [optimized, +5/-4 matrix (5:-4)] ktup: 6
join: 63, opt: 48, gap-pen: -16/-4, width: 16

Scan time: 0.116

The best scores are: opt

gi|27735142|ref|NM_173674.1| Homo sapiens disc (2010) [f] 7311
gi|27735142|ref|NM_173674.1| Homo sapiens disc (2010) [r] 67

>>gi|27735142|ref|NM_173674.1| Homo sapiens discoidin, C (2010 nt)
initn: 7311 init1: 7311 opt: 7311
99.932% identity in 1464 nt overlap (1-1464:282-1745)

LEX				10	20	30
				ATGACATCTAAGAATTATCCCGGGACCTAC		
					
gi 277	CACCTAGTGACTTATCAGGATAGTGGCACAATGACATCTAAGAATTATCCCGGGACCTAC					
	260	270	280	290	300	310
LEX		40	50	60	70	80
		CCCAATCACACTGTTTGC	CGAAAAGACAATTACAGTACCAAAGGGGAAAAGACTGATTCTG			
					
gi 277	CCCAATCACACTGTTTGC	CGAAAAGACAATTACAGTACCAAAGGGGAAAAGACTGATTCTG				
	320	330	340	350	360	370
LEX		100	110	120	130	140
		AGGTTGGGAGATTG	GATATCGAATCCAGACCTGTGCTTCTGACTATCTTCTCTTCACC			
					
gi 277	AGGTTGGGAGATTG	GATATCGAATCCAGACCTGTGCTTCTGACTATCTTCTCTTCACC				
	380	390	400	410	420	430
LEX		160	170	180	190	200
		AGCTCTTCAGATCAATATGGTCCATACTGTGGAAGTATGACTGTTCCCAAAGAACTCTTG				
					
gi 277	AGCTCTTCAGATCAATATGGTCCATACTGTGGAAGTATGACTGTTCCCAAAGAACTCTTG					
	440	450	460	470	480	490
LEX		220	230	240	250	260
		TTGAACACAAGTGAAGTAACCGTCCGCTTTGAGAGTGGATCCCACATTTCTGGCCGGGGT				
					
gi 277	TTGAACACAAGTGAAGTAACCGTCCGCTTTGAGAGTGGATCCCACATTTCTGGCCGGGGT					
	500	510	520	530	540	550
LEX		280	290	300	310	320
		TTTTTGCTGACCTATGCGAGCAGCGACCATCCAGATTTAATAACATGTTTGGAACGAGCT				
					
gi 277	TTTTTGCTGACCTATGCGAGCAGCGACCATCCAGATTTAATAACATGTTTGGAACGAGCT					
	560	570	580	590	600	610
	340	350	360	370	380	390

```
LEX      AGCCATTATTTGAAGACAGAATACAGCAAATTCTGCCAGCTGGTTGTAGAGACGTAGCA
          .....
gi|277   AGCCATTATTTGAAGACAGAATACAGCAAATTCTGCCAGCTGGTTGTAGAGACGTAGCA
          620      630      640      650      660      670

          400      410      420      430      440      450
LEX      GGAGACATTTCTGGGAATATGGTAGATGGATATAGAGATACCTCTTTATTGTGCAAAGCT
          .....
gi|277   GGAGACATTTCTGGGAATATGGTAGATGGATATAGAGATACCTCTTTATTGTGCAAAGCT
          680      690      700      710      720      730

          460      470      480      490      500      510
LEX      GCCATCCATGCAGGAATAATTGCTGATGAACTAGGTGGCCAGATCAGTGTGCTTCAGCGC
          .....
gi|277   GCCATCCATGCAGGAATAATTGCTGATGAACTAGGTGGCCAGATCAGTGTGCTTCAGCGC
          740      750      760      770      780      790

          520      530      540      550      560      570
LEX      AAAGGGATCAGTCGATATGAAGGGATTCTGGCCAATGGTGTTCCTTCGAGGGATGGTTCC
          .....
gi|277   AAAGGGATCAGTCGATATGAAGGGATTCTGGCCAATGGTGTTCCTTCGAGGGATGGTTCC
          800      810      820      830      840      850

          580      590      600      610      620      630
LEX      CTGTCAGACAAGCGATTTCTGTTTACCTCCAATGGTTGCAGCAGATCCTTGAGTTTTGAA
          .....
gi|277   CTGTCAGACAAGCGATTTCTGTTTACCTCCAATGGTTGCAGCAGATCCTTGAGTTTTGAA
          860      870      880      890      900      910

          640      650      660      670      680      690
LEX      CCTGACGGGCAAATCAGAGCTTCTTCCTCATGGCAGTCGGTCAATGAGAGTGGAGACCAA
          .....
gi|277   CCTGACGGGCAAATCAGAGCTTCTTCCTCATGGCAGTCGGTCAATGAGAGTGGAGACCAA
          920      930      940      950      960      970

          700      710      720      730      740      750
LEX      GTTCACTGGTCTCCTGGCCAAGCCCGACTTCAGGACCAAGGCCCATCATGGGCTTCGGGC
          .....
gi|277   GTTCACTGGTCTCCTGGCCAAGCCCGACTTCAGGACCAAGGCCCATCATGGGCTTCGGGC
          980      990      1000      1010      1020      1030

          760      770      780      790      800      810
LEX      GACAGTAGCAACAACCACAAACCACGAGAGTGGCTGGAGATCGATTTGGGGGAGAAAAAG
          .....
gi|277   GACAGTAGCAACAACCACAAACCACGAGAGTGGCTGGAGATCGATTTGGGGGAGAAAAAG
          1040      1050      1060      1070      1080      1090

          820      830      840      850      860      870
LEX      AAAATAACAGGAATTAGGACCACAGGATCTACACAGTCGAACTTCAACTTTTATGTTAAG
          .....
gi|277   AAAATAACAGGAATTAGGACCACAGGATCTACACAGTCGAACTTCAACTTTTATGTTAAG
          1100      1110      1120      1130      1140      1150

          880      890      900      910      920      930
LEX      AGTTTTGTGATGAACTTCAAAAACAATAATTCTAAGTGAAGACCTATAAAGGAATTGTG
          .....
gi|277   AGTTTTGTGATGAACTTCAAAAACAATAATTCTAAGTGAAGACCTATAAAGGAATTGTG
          1160      1170      1180      1190      1200      1210

          940      950      960      970      980      990
```

```

LEX      AATAATGAAGAAAAGGTGTTTCAGGGTAACTCTAACTTTCGGGACCCAGTGCAAAACAAT
          .....
gi|277   AATAATGAAGAAAAGGTGTTTCAGGGTAACTCTAACTTTCGGGACCCAGTGCAAAACAAT
          1220      1230      1240      1250      1260      1270

          1000      1010      1020      1030      1040      1050
LEX      TTCATCCCTCCCATCGTGGCCAGATATGTGCGGGTTGTCCCCCAGACATGGCACCAGAGG
          .....
gi|277   TTCATCCCTCCCATCGTGGCCAGATATGTGCGGGTTGTCCCCCAGACATGGCACCAGAGG
          1280      1290      1300      1310      1320      1330

          1060      1070      1080      1090      1100      1110
LEX      ATAGCCTTGAAGGTGGAGCTCATTGGTTGCCAGATTACACAAGGTAATGATTCATTGGTG
          .....
gi|277   ATAGCCTTGAAGGTGGAGCTCATTGGTTGCCAGATTACACAAGGTAATGATTCATTGGTG
          1340      1350      1360      1370      1380      1390

          1120      1130      1140      1150      1160      1170
LEX      TGGCGCAAGACAAGTCAAAGCACCAGTGTTTCAACTAAGAAAGAAGATGAGACAATCACA
          .....
gi|277   TGGCGCAAGACAAGTCAAAGCACCAGTGTTTCAACTAAGAAAGAAGATGAGACAATCACA
          1400      1410      1420      1430      1440      1450

          1180      1190      1200      1210      1220      1230
LEX      AGGCCCATCCCCTCGGAAGAAACATCCACAGGAATAAACATTACAACGGTGGCTATTCCA
          .....
gi|277   AGGCCCATCCCCTCGGAAGAAACATCCACAGGAATAAACATTACAACGGTGGCTATTCCA
          1460      1470      1480      1490      1500      1510

          1240      1250      1260      1270      1280      1290
LEX      TTGGTGCTCCTTGTTGTCCTGGTGTTTGCTGGAATGGGGATCTTTGCAGCCTTTAGAAAG
          .....
gi|277   TTGGTGCTCCTTGTTGTCCTGGTGTTTGCTGGAATGGGGATCTTTGCAGCCTTTAGAAAG
          1520      1530      1540      1550      1560      1570

          1300      1310      1320      1330      1340      1350
LEX      AAGAAGAAGAAAGGAAGTCCGTATGGATCAGCGGAGGCTCAGAAAACAGACTGTTGGAAG
          .....
gi|277   AAGAAGAAGAAAGGAAGTCCGTATGGATCAGCAGAGGCTCAGAAAACAGACTGTTGGAAG
          1580      1590      1600      1610      1620      1630

          1360      1370      1380      1390      1400      1410
LEX      CAGATTAAATATCCCTTTGCCAGACATCAGTCAGCTGAGTTTACCATCAGCTATGATAAT
          .....
gi|277   CAGATTAAATATCCCTTTGCCAGACATCAGTCAGCTGAGTTTACCATCAGCTATGATAAT
          1640      1650      1660      1670      1680      1690

          1420      1430      1440      1450      1460
LEX      GAGAAGGAGATGACACAAAAGTTAGATCTCATCACAAGTGATATGGCAGGTAA
          .....
gi|277   GAGAAGGAGATGACACAAAAGTTAGATCTCATCACAAGTGATATGGCAGGTAACTCCGT
          1700      1710      1720      1730      1740      1750

gi|277   TGA CTGCCAAAATAGCATCCCCAACGTGCAGCCCTCCGCATCTATCAGCAGGTTGCCCCG
          1760      1770      1780      1790      1800      1810

>>gi|27735142|ref|NM_173674.1| Homo sapiens discoidin, C (2010 nt)
rev-comp initn: 57 initl: 57 opt: 67
61.364% identity in 44 nt overlap (492-449:1780-1823)

```

```
      520      510      500      490      480      470
LEX-  ACTGATCCCTTTGCGCTGAAGCACACTGATCTGGCCACCTAGTTCATCAGCAATTATTCC
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|277 GTTGACTGCCAAAATAGCATCCCCAACGTGCAGCCCTCCGCATCTATCAGCAGGTTGCCC
      1750      1760      1770      1780      1790      1800

      460      450      440      430      420      410
LEX-  TGCATGGATGGCAGCTTTGCACAATAAAGAGGTATCTCTATATCCATCTACCATATTCCC
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|277 CGGATGGATCTCAGAGATGAGGATCGGAACACCATGTTCTTTCCCACCCTAACAACAACA
      1810      1820      1830      1840      1850      1860
```

1464 residues in 1 query sequences

2010 residues in 1 library sequences

Scomplib [version 3.3t05 March 30, 2000]

start: Mon Oct 6 13:26:41 2003 done: Mon Oct 6 13:26:41 2003

Scan time: 0.116 Display time: 0.100

Function used was FASTA